IEEE Std 1609.2-2016 Guidance Note 1: ECDSA

William Whyte, Security Innovation
Version 0.2, 2016-08-03

1 Introduction

This note identifies an area in IEEE Std 1609.2-2016 that has been reported as potentially misleading or ambiguous and gives guidance on a recommended interpretation of the standard.

The potentially misleading or ambiguous text concerns the encoding of signatures for the Elliptic Curve Digital Signature Algorithm (ECDSA).

This guidance note is structured as follows. Section 2 provides a summary of the issue. Section 3 provides the recommended interpretation. Section 3.1 provides the full text from the relevant sections of IEEE Std 1609.2-2016. Section 3.2 provides full replacement text that should be considered for incorporation into a revision or amendment of 1609.2.

2 Summary of issue

When an ECDSA signature is generated over the NIST p256 curve, there are two large prime numbers that are used in the calculations:

- p = 11 5792089210356248 762697446949407 573530086143415 290314195533631 308867097853951 the prime number used to define all operations on the curve; each curve point is two numbers (x, y) which are less than or equal to p.
- n = 11 57920 89210356248 762697446949407 573529996955224 135760342422259 061068512044369 the "order of the generator" on the curve, informally an upper limit on the number of operations that can be carried out before results start to repeat themselves. NOTE that n is less than p.

There are two different "flavors" of ECDSA:

- Standard, as per Federal Information Processing Standard (FIPS) 186-4. In this case, the output of signing is two integers (r, s) which are both less than n.
- Fast-verification enabled, as per Standards for Efficient Cryptography (SEC) #1. In this case, the output of signing is an elliptic curve point R and an integer s, with s less than n.

The EcdsaP256Signature data structure in IEEE 1609.2 supports both flavors of ECDSA. It consists of an elliptic curve point r and an octet string s.

- For standard ECDSA, the integer r calculated in FIPS 186-4 is encoded in the x-coordinate of the point r; the integer s calculated in FIPS 186-4 is encoded in the octet string s.
- For ECDSA with fast verification, the elliptic curve point r calculated in SEC1 is encoded as the point r; the integer s calculated in SEC1, which is the same as the s calculated in FIPS 186-4, is encoded in the octet string s.

The potential area of confusion affects standard ECDSA signatures.

One possible area of confusion arises from the fact that r is used to indicate **both** the output r value from signing (with either FIPS 186-4 or SEC 1) **and** the field in the 1609.2 data structure that encodes the r-value. To reduce the possibility of this confusion, we will use "rSig" below to denote the field in 1609.2, and "r" to denote the value output from signing.

It happens to be the case that for standard ECDSA signatures, r is calculated as follows:

- 1. Generate a random value *k* mod *n*.
- 2. Generate a point R = k*G, G the "generator"
- 3. Set r' equal to the x-coordinate of R, i.e. an integer mod p
- 4. Set r equal to r' mod n.

Although 1609.2 is written to make it clear that this final r value is what is encoded in the x-coordinate of the point rSig in the EcdsaP256Signature structure, an implementer who knows a bit about ECDSA might be confused and think that since r is encoded in the x-coordinate of an elliptic curve point, therefore r is intended to be the x-value r' that is calculated internally to the ECDSA process at step 3 above. This is not the case. When the signature is calculated per FIPS 186-4, the r-value encoded in the rSig field of the EcdsaP256Signature is the r-value output from FIPS 186-4 signing. The r-value encoded in the rSig field of the EcdsaP256Signature is NOT the intermediate r' value that is created, but not output, during the signing process.

3 Recommended interpretation

The recommended interpretation is to follow the straightforward meaning of the text in IEEE Std 1609.2-2016 and, for a standard ECDSA signature, encode r as the x-coordinate of the point, rSig, on curve NISTp256 even though r is actually an integer mod n.

3.1 Original text

6.3.30 EcdsaP256Signature

This structure represents an ECDSA signature. The signature is generated as specified in 5.3.1.

If the signature process followed the specification of FIPS 186-4 and output the integer r, r is represented as an EccP256CurvePoint indicating the selection x-only.

If the signature process followed the specification of SEC 1 and output the elliptic curve point *R* to allow for fast verification, *R* is represented as an EccP256CurvePoint indicating the choice compressed-y-0, compressed-y-1, or uncompressed at the sender's discretion.

3.2 Proposed replacement text

6.3.30 EcdsaP256Signature

This structure represents an ECDSA signature. The signature is generated as specified in 5.3.1.

If the signature process followed the specification of FIPS 186-4 and output an integer r, r is represented as the x-coordinate of the EccP256CurvePoint rSig, with the CHOICE of type in rSig set to x-only. <add footnote>

If the signature process followed the specification of SEC 1 and output the elliptic curve point R to allow for fast verification, R is represented as the EccP256CurvePoint rSig, with the CHOICE of type in rSig set to compressed-y-0, compressed-y-1, or uncompressed at the sender's discretion.

For both the FIPS 186-4 and the SEC 1 approaches, the integer *s* output from signing is encoded in sSig as a 32-byte integer, encoded in network byte order and left-padded with zeroes if necessary.

Footnote: In this case, the x-coordinate of the point rsig contains the value r *output* from FIPS 186-4 signing, not any elliptic curve point x-value that is calculated as an intermediate value within FIPS 186-4 signing.